Available online at www.sciencedirect.com

ScienceDirect

Journal of Computational and Applied Mathematics 199 (2007) 238–244

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

An efficient estimate based on FFT in topological verification method

Yasuaki Hiraoka*, Toshiyuki Ogawa

Department of Mathematical Science, Graduate School of Engineering Science, Osaka University, Japan

Received 14 December 2004

Abstract

In this paper, localized patterns of the quintic Swift–Hohenberg equation are studied. A numerical verification method with the Conley index theory developed in Zgliczyński and Mischaikow [Rigorous numerics for partial differential equations: the Kuramoto–Sivashinsky equation, Found. Comput. Math. 1 (2001) 255–288] is used in order to prove these patterns. A new technique to efficiently obtain estimates for nonlinear terms is presented. The key idea is based on the pseudo-spectral method. It is shown that this technique is inevitable for the verification of the localized patterns.

Keywords: Numerical verification; Conley index; FFT

1. Introduction

In this paper, we deal with the quintic Swift–Hohenberg equation:

$$u_t = \left\{ \nu - \left(1 + \frac{\partial^2}{\partial x^2} \right)^2 \right\} u + \mu u^3 - u^5 \tag{1}$$

under the periodic boundary condition $u(x + L_0, t) = u(x, t)$, where $x, t, u(x, t) \in \mathbb{R}$. In particular, our main interest is devoted to stationary localized patterns of this equation. Sakaguchi and Brand [10] observe by computer simulations that the quintic Swift–Hohenberg equation may have many types of stable localized stationary solutions in suitable parameter regions. Furthermore, they heuristically explain the relation between the existence of the stable localized patterns and the coexistence of stable uniform solutions and stable spatially periodic (roll) solutions, which is realized in the subcritical region ($\nu < 0$). Let us comment that the original Swift–Hohenberg equation corresponds to (1) with $\mu = -1$ and no quintic term. In this case, the equation describes the onset of Rayleigh–Bénard heat convection and the parameter ν represents the Rayleigh number [12].

Let us observe the following bifurcation diagram shown in Fig. 1. This diagram represents the bifurcation branches for approximate stationary solutions of (1) with $\mu = 3$. These bifurcation branches are numerically calculated by the pseudo arclength method [7]. From the figure, we can observe that the pure mode branch bifurcates from the trivial solution

* Corresponding author.
E-mail address: hiraoka@nsc.es.hokudai.ac.jp (Y. Hiraoka).

0377-0427/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2005.08.036

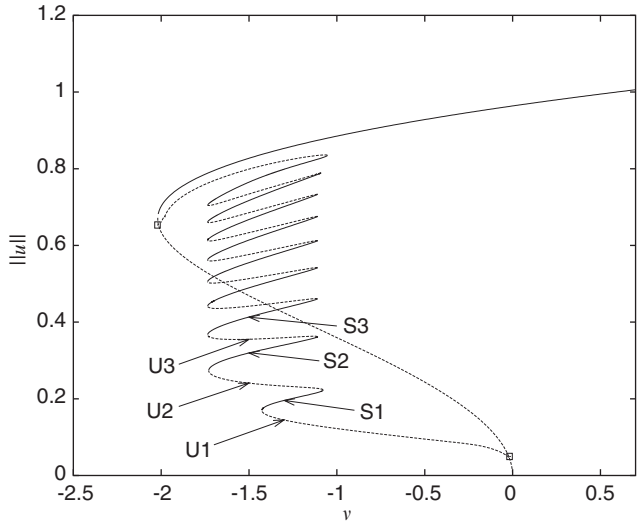

Fig. 1. Bifurcation diagram of (1) with $\mu = 3$.

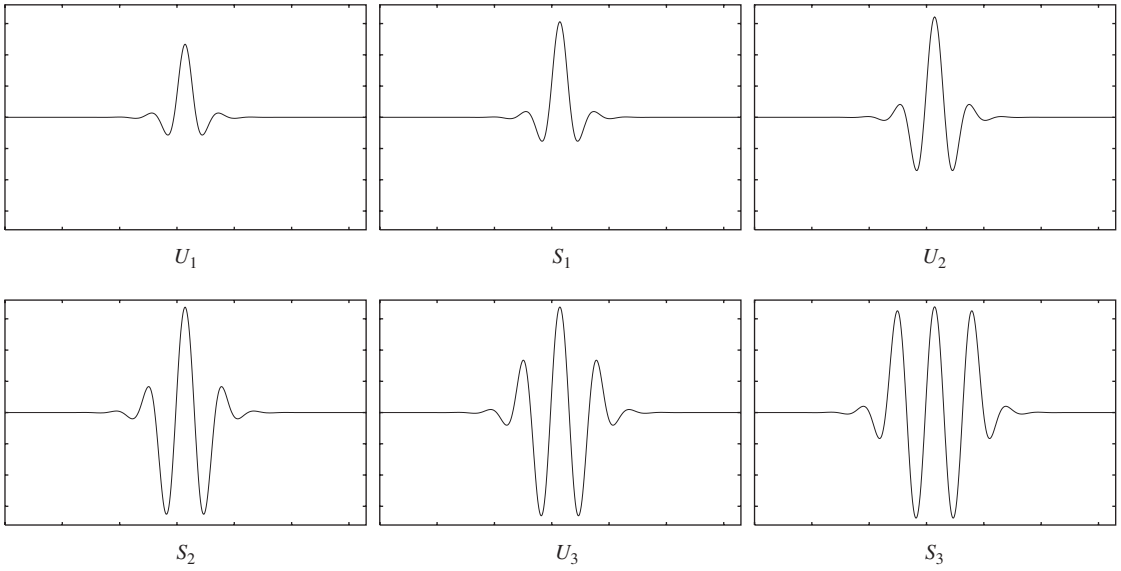

Fig. 2. Profiles of equilibria on the segments U_k, S_k, $k = 1, 2, 3$ ($v = -1.3$).

$u(x, t) = 0$ around $v \approx 0$. This bifurcation occurs as a subcritical pitchfork type. Then, right after the bifurcation from the trivial solution, a mixed mode solution bifurcates from this pure mode branch. The bifurcation structure around the trivial solution can be studied by using weak nonlinear analysis based on the center manifold reduction. These analytical results are consistent with the bifurcation structure shown in Fig. 1 (see [5]).

However, let us note the mixed mode solutions in a parameter region far away from $v = 0$. We can observe that the saddle-node bifurcations repeat on the mixed mode branch between $v \approx -1.735$ and -1.110. This behavior cannot be predictable from the weak nonlinear analysis. We call such a bifurcation branch "a snaky bifurcation branch" in this paper and study it in detail. Let us denote each segment on the snaky bifurcation branch between two successive saddle-node bifurcations by U_k and S_k as is described in the figure. Fig. 2 shows the wave profiles on the lower segments. We can observe localized patterns which correspond to the numerical results shown in [10].

Since the bifurcation branches in Fig. 1 are approximate, we cannot insist whether these localized solutions on the snaky bifurcation branch really exist or not. Hence we study the existence of these localized solutions from the viewpoint of the rigorous numerics. We adopt a topological verification method developed in the paper [14] for the verification. In the topological verification method, we deal with the Fourier expansion to derive the infinite dimensional dynamical systems. However, our targets are localized patterns and many Fourier modes are required to express these patterns. This causes the trouble of the computational cost for the verification. In this paper, we present an efficient improvement of the topological verification method to overcome the difficulty. This technique plays an crucial role for the verification of the localized patterns.

Let us finally remark that there have been several works [1,6,13] which deal with the stationary problem of (1), i.e. the fourth-order ordinary differential equation. In these works, the localized solutions are analyzed from the viewpoint of the homoclinic orbits for the trivial solution by the asymptotic expansion. They prove the existence of the homoclinic orbits around the first bifurcation point. However, it seems to be difficult by these approaches to relate the existence of the homoclinic orbits with the snaky branch. In the paper, we directly study the stationary solutions on the snaky bifurcation branches by numerical verification.

2. Topological verification method

At first, we recall the elementary results of the Conley index theory, which plays an important role throughout this paper (see [2,11] for more detailed introductions). Let $\varphi : \mathbb{R}^+ \times X \to X$ be a semiflow on a locally compact metric space X, where $\mathbb{R}^+ = [0, \infty)$. A complete orbit through x is a function $\gamma_x : \mathbb{R} \to X$ such that $\gamma_x(0) = x$ and $\varphi(t, \gamma_x(s)) = \gamma_x(t+s)$ for any $s \in \mathbb{R}$ and $t \geqslant 0$. Given a subset $N \subset X$, $\mathrm{Inv}(N, \varphi) := \{x \in N \mid \exists \text{a complete orbit } \gamma_x : \mathbb{R} \to N\}$ is called the maximal invariant set in N. A compact set N is defined as an isolating neighborhood if its maximal invariant set is included in the interior of N, i.e. $\mathrm{Inv}(N, \varphi) \subset \mathrm{Int}(N)$, where $\mathrm{Int}(N)$ denotes the interior of N. This maximal invariant set is called the isolated invariant set.

Definition 1. A pair of compact sets (N, N^+) is called an index pair for the isolated invariant set S if $N^+ \subset N$ and the following three conditions hold.

1. $\mathrm{Cl}(N \backslash N^+)$ is an isolating neighborhood of S.
2. N^+ is positively invariant in N.
3. If $x \in N$ and $\varphi(\mathbb{R}^+, x) \not\subset N$, then there exists a $t \geqslant 0$ such that $\varphi([0, t], x) \subset N$ and $\varphi(t, x) \in N^+$.

Here $\mathrm{Cl}(N \backslash N^+)$ denotes the closure of $N \backslash N^+$. The above condition 2 is equivalent to say that if $x \in N^+$ and $\varphi([0, t], x) \subset N$ for some $t > 0$, then $\varphi([0, t], x) \subset N^+$. The property 3 means that every orbit which leaves N in forward time has to go through N^+ before leaving N.

Definition 2. Let (N, N^+) be an index pair for an isolated invariant set S. The Conley index of S is defined by the relative homology group of (N, N^+):

$$CH_*(S) := H_*(N, N^+).$$

It should be noted that the above definition is well-defined [2,11]. More precisely, we can show that there exists an index pair for any given isolated invariant set. In addition, if (N_1, N_1^+) and (N_2, N_2^+) are two different index pairs for the same isolated invariant set, then $CH_*(\mathrm{Inv}(\mathrm{Cl}(N_1 \backslash N_1^+), \varphi)) \cong CH_*(\mathrm{Inv}(\mathrm{Cl}(N_2 \backslash N_2^+), \varphi))$. Hence, given an index pair (N, N^+) for S, we also use the notation $CH_*(N, N^+)$ to express the Conley index of S.

Next, following papers [3,14], we discuss how index information can be used to verify the existence of the stationary solution. By using the Fourier cosine expansion $u(x, t) = \sum_{j \in \mathbb{Z}} a_j(t) \cos(j k_0 x)$, where $k_0 = 2\pi / L_0$ and $a_{-j} = a_j$, the system of ODEs for (1) is given by

$$\dot{a}_j = f_j(a) = \zeta_j a_j + \mu f_j^{(3)}(a) - f_j^{(5)}(a), \quad j = 0, 1, \ldots . \tag{2}$$

Here $\dot{a}_j$ represents the time derivative of a_j, $\zeta_j = \nu - (1 - j^2k_0^2)^2$ and

$$f_j^{(3)}(a) = \sum_{\substack{m_1+m_2+m_3=j \\ m_i\in\mathbb{Z}}} a_{m_1}a_{m_2}a_{m_3}, \quad f_j^{(5)}(a) = \sum_{\substack{m_1+m_2+m_3+m_4+m_5=j \\ m_i\in\mathbb{Z}}} a_{m_1}a_{m_2}a_{m_3}a_{m_4}a_{m_5}.$$

By the general theory [8], this system of ODEs is well-posed on the space

$$\mathscr{X} := \left\{ a = (a_j) \mid \|a\|_{\mathscr{X}}^2 := \sum_{j\geqslant 0} a_j^2(1 + k_0^2 j^2)^4 < \infty \right\}.$$

That is to say the system of ODEs (2) defines a continuous semiflow $\varphi : \mathbb{R}^+ \times \mathscr{X} \to \mathscr{X}$. On this setting, a stationary solution can be regarded as an equilibrium point of $f_j(a) = 0$, $j = 0, 1, \ldots$.

In order to apply the Conley index theory to (2) we introduce a subset $Y \subset \mathscr{X}$ by

$$Y = \prod_{j\geqslant 0} Y_j, \quad Y_j := \begin{cases} \mathbb{R}, & j = 0, 1, \ldots, m, \\ \left[-\dfrac{c}{j^s}, \dfrac{c}{j^s}\right], & j > m, \end{cases} \tag{3}$$

where $s > \frac{9}{2}$ and $c > 0$. Moreover, let us consider the product topology on Y. Then it is easy to check that the topology induced by the metric $\|\cdot\|_{\mathscr{X}}$ and the product topology are equivalent on the topological space Y. Moreover, by the standard energy estimates, we also show that there exists a compact positively invariant set X in Y which contains a global attractor of (2). Obviously, the semiflow φ induces a new semiflow $\bar{\varphi} : \mathbb{R}^+ \times X \to X$ on the positively invariant set and the arguments throughout the paper are based on this setting. We refer to [4] for details.

In order to reduce the infinite dimensional problem to finite dimensional one, it is convenient to use the following notation:

$$a = (a_F, a_I), \quad a_F = (a_0, a_1, \ldots, a_m), \quad a_I = (a_{m+1}, a_{m+2}, \ldots),$$
$$f(a) = (f_F(a), f_I(a)), \quad f_F(a) = (f_0(a), f_1(a), \ldots, f_m(a)), \quad f_I = (f_{m+1}(a), f_{m+2}(a), \ldots).$$

In the following, subscripts F and I represent the finite part and the infinite part, respectively. Let us remark that we choose m in such a way that the essential dynamics of (2) should be contained in the finite part.

Let $g_F(a_F) := f_F(a_F, a_I = 0)$ be the Galerkin approximation of $f(a)$ and $\bar{a} = (\bar{a}_F, 0)$ be an approximate equilibrium point such that $g_F(\bar{a}_F) \approx 0$. For the numerical verification, it is necessary to prepare such an approximate solution since we try to verify the existence of a stationary solution around the approximate solution. Note that the error term of the Galerkin approximation is $r(a_F, a_I) := f_F(a_F, a_I) - g_F(a_F)$.

Given a compact set, we need to check the vector field on the boundary for the computation of the Conley index. For this purpose, it may be convenient to introduce a new variable b_j, $j = 0, 1, \ldots$, by

$$(Pb_F + \bar{a}_F, b_I) = (a_F, a_I), \tag{4}$$

where the eigenvectors p_j, $j = 0, 1, \ldots, m$, of the Jacobi matrix $Dg_F(\bar{a}_F)$ are taken to be column vectors for the matrix P. We assume that $Dg_F(\bar{a}_F)$ is diagonalizable and the eigenvalues λ_j, $j = 0, 1, \ldots, m$, corresponding to the eigenvectors p_j satisfy $Re(\lambda_j) \neq 0$. This transformation defined by (4) is denoted by $T : (b_F, b_I) \mapsto (a_F, a_I)$.

After the Taylor expansion of $g_F(a_F)$ at $\bar{a}_F$ with the new variable b_j, the original dynamical system can be represented by

$$\dot{b}_j = h_j(b) := \lambda_j b_j + R_j(b), \quad j = 0, 1, \ldots, \tag{5}$$

where $\lambda_j := \zeta_j$, $R_j(b) := \mu f_j^{(3)}(Pb_F + \bar{a}_F, b_I) - f_j^{(5)}(Pb_F + \bar{a}_F, b_I)$ for the infinite part ($j > m$) and

$$R_F(b) = P^{-1}(g_F(\bar{a}_F) + \tfrac{1}{2}D^2 g_F(\bar{a}_F)(Pb_F)^2 + \cdots + \tfrac{1}{5!}D^5 g_F(\bar{a}_F)(Pb_F)^5 + r(b_F, b_I)). \tag{6}$$

Let us consider a compact set

$$W = W_F \times W_I, \quad W_F := \prod_{j=0}^{m} [b_j^-, b_j^+], \quad W_I := \prod_{j>m} \left[-\frac{c}{j^s}, \frac{c}{j^s} \right], \tag{7}$$

where c and s are positive constants. For this compact set, we call $F_j^- := \{b \in \partial W \mid b_j = b_j^-\}$ and $F_j^+ := \{b \in \partial W \mid b_j = b_j^+\}$ the faces of W. Then the transversality of the vector field (5) on the faces $F_j^\pm$ is checked by $h_j(b)|_{b \in F_j^\pm} \neq 0$.

Definition 3. A compact set $W = \prod_{j \geqslant 0} [b_j^-, b_j^+] \ni 0$ given by the form (7) is defined as a lifting set for the dynamical system (5) if the following conditions are satisfied.

1. The vector field is transverse on all the faces of W.
2. The vector field on the boundary $W_F \times \partial W_I$ is inward to W, i.e., $h_j(b)|_{b \in F_j^\pm} \lesseqgtr 0$ for $j > m$.

Let us define $W_F^+ \subset \partial W_F$ by the union of the faces of W_F where the vector field is directed outward. We also define $W^+ := W_F^+ \times W_I$. From the definition of the lifting set, (W, W^+) can be an index pair. Moreover, we can check $CH_*(W, W^+) = H_*(W, W^+) = H_*(W_F, W_F^+)$. Therefore the Conley index of the lifting set is essentially determined by its finite part. From this argument, we often use the notation $CH_*(W_F, W_F^+)$ for the Conley index of the index pair (W, W^+).

Now we are ready to show a theorem which plays a central role for the topological verification method of stationary solutions.

Theorem 4 (*Zgliczyński and Mischaikow [14]*). *Let W be a lifting set for the dynamical system* (5). *If the Conley index of the lifting set W satisfies*

$$CH_j(W_F, W_F^+) \cong \begin{cases} \mathbb{Z}_2, & j = k, \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

for some $k \in \{0, 1, \ldots, m\}$, then there exists an equilibrium point of (2) *in $T \cdot W$.*

Let us remark that, from the viewpoint of the rigorous numerics, it is necessary to explicitly construct a lifting set which satisfies the condition in Theorem 4. It is known that, from the argument in [3,14], the form of the lifting set (7) enables us to obtain the estimates for the vector field $\{h_j(b)\}$ in W by using the interval arithmetic. Therefore we can rigorously check the sufficient condition in Theorem 4.

3. Efficient method for estimates of error bounds

In this section, we discuss an efficient method to obtain estimates of nonlinear terms. From the error bounds studied in [3], we have to rigorously calculate the finite sum given by

$$\sum_{\substack{m_1+m_2+\cdots+m_p=j \\ |m_i| \leqslant m}} a_{m_1} a_{m_2} \cdots a_{m_p}, \quad j = 0, 1, \ldots, m \tag{9}$$

for the estimates of the pth nonlinear terms. The direct calculation of this collection requires the $O(m^p)$ computational cost. Obviously, it is not efficient for large m and p. We present an improvement to efficiently calculate the finite sums (9). The key idea comes from the pseudo spectral method, which is well-known as one of the numerical simulation methods for studying nonlinear PDEs.

We discuss the quadratic nonlinearity ($p=2$) for the sake of simplicity. Let us consider the discrete Fourier transform and its inverse:

$$a_l = \mathscr{F}(u)|_l = \sum_{j=0}^{2m-1} u(x_j)\mathrm{e}^{-\mathrm{i}lk_0x_j}, \tag{10}$$

$$u(x_j) = \mathscr{F}^{-1}(a)|_j = \frac{1}{2m}\sum_{l=-m+1}^{m} a_l\mathrm{e}^{\mathrm{i}lk_0x_j}. \tag{11}$$

Here $\{x_j = (L_0/2m)j\}$, $j = 0, 1, \ldots, 2m-1$, are grid points in the interval $[0, L_0]$ and $k_0 = 2\pi/L_0$.

The basic idea of the pseudo-spectral method is the following. First of all, we pull back the Fourier coefficients $\{a_l\}$ to the original variable $\{u(x_j)\}$ by (11). Then we calculate the nonlinear term $\{u^2(x_j)\}$ at each grid point. Finally, the sum (9) may be calculated by transforming $\{u^2(x_j)\}$ to each Fourier mode by (10). These arguments can be described as follows.

$$\begin{aligned} c_l &= \sum_{j=0}^{2m-1} u(x_j)^2\mathrm{e}^{-\mathrm{i}lk_0x_j} \\ &= \frac{1}{2m}\sum_{\substack{m_1+m_2=l \\ -m+1\leqslant m_i\leqslant m}} a_{m_1}a_{m_2} + \frac{1}{2m}\sum_{\substack{m_1+m_2=l\pm 2m \\ -m+1\leqslant m_i\leqslant m}} a_{m_1}a_{m_2}. \end{aligned} \tag{12}$$

The second term of (12) is called an aliasing error. This error is incorporated into the convolution because two different Fourier modes by modulo $2m$ cannot be distinguished due to the discretization of the space. One of the methods to remove aliasing errors is as follows [9]. We extend the size of the discrete Fourier transform from $2m$ to $2m\delta$ for $\delta > 1$. In addition, let $\{a_j\}$, $j = -\delta m + 1, \ldots, \delta m$, be taken as the same values of the original Fourier coefficients for $|j| \leqslant m$ and $a_j = 0$ for the remainder. Then, the same calculation shown in (12) for the extended Fourier coefficients leads to

$$\hat{c}_l = \frac{1}{2m\delta}\sum_{\substack{m_1+m_2=l \\ |m_i|\leqslant m}} a_{m_1}a_{m_2} + \frac{1}{2m\delta}\sum_{\substack{m_1+m_2=l\pm 2\delta m \\ |m_i|\leqslant m}} a_{m_1}a_{m_2}.$$

Hence, if $\delta > \frac{3}{2}$, then the second term which causes the aliasing error can be completely eliminated. Finally the finite sum (9) for $p = 2$ is calculated by $2m\delta\hat{c}_l$.

The same approach can be applied to the general nonlinear term (9) by taking δ suitably. For example, the similar calculation leads to the following lemma, which is needed for the verification in the quintic Swift–Hohenberg equation.

Lemma 5. *The constant δ to remove aliasing errors should satisfy $\delta > 2$ for $p = 3$ and $\delta > 3$ for $p = 5$, respectively.*

Note that by using the fast Fourier transform (FFT) equipped with the interval arithmetic we obtain the rigorous error bounds of (9) quite efficiently. The computational cost only depends on the calculation of FFT and reduces to the order $m \log m$. This technique becomes indispensable for topological verifications of the localized patterns of the quintic Swift–Hohenberg equation.

Let us denote an approximate solution on the snaky branch in Fig. 1 by $u(x; v, l) = \sum_{|j|\leqslant m} a_j \cos(jk_0x)$, where the Fourier coefficients $\{a_j\}$ correspond to the approximate equilibrium point on each segment $l = U_k, S_k$ at the parameter value v. Then, we obtain the following theorem by the topological verification method with FFT algorithm.

Theorem 6. *Let $v = -1.3$, $k_0 = 0.1$, and $\mu = 3$. Then there exists a stationary solution $u_*(x; v, l)$ of the quintic Swift–Hohenberg equation in the neighborhood of each approximate solution $u(x; v, l)$, $l = U_k, S_k$, $k = 1, 2, 3$, such that*

$$\|u_*(\cdot; v, U_1) - u(\cdot; v, U_1)\|_{L^2} \leqslant 1.04077019 \times 10^{-8},$$

$$\|u_*(\cdot; v, S_1) - u(\cdot; v, S_1)\|_{L^2} \leqslant 1.57739803 \times 10^{-8},$$

$$\|u_*(\cdot; v, U_2) - u(\cdot; v, U_2)\|_{L^2} \leqslant 2.44819377 \times 10^{-8},$$

$$\|u_*(\cdot; v, S_2) - u(\cdot; v, S_2)\|_{L^2} \leqslant 4.31155312 \times 10^{-8},$$

$$\|u_*(\cdot; v, U_3) - u(\cdot; v, U_3)\|_{L^2} \leqslant 2.83246161 \times 10^{-9},$$

$$\|u_*(\cdot; v, S_3) - u(\cdot; v, S_3)\|_{L^2} \leqslant 7.47772691 \times 10^{-9}.$$

Here, we set $c = 1.0$ and $s = 5$ for the power decay property (7) in the verifications. Moreover, the dimension for the finite part is chosen as $m = 256$ for U_k, S_k, $k = 1, 2$, and $m = 512$ for U_3, S_3. From the above numerical results, we can expect that there really exist a snaky bifurcation branch in the quintic Swift–Hohenberg equation (1) as is shown in Fig. 1. In the paper [5], not only the existence of the localized stationary solutions but also the bifurcation structure of the quintic Swift–Hohenberg equation is discussed in detail.

Acknowledgements

The work of the first author was partially supported by Grant-in-Aid for J.S.P.S. Fellows, 03948.

References

[1] C.J. Budd, G.W. Hunt, R. Kuske, Asymptotics of cellular buckling close to the Maxwell load, Proc. Roy. Soc. London Ser. A 457 (2001) 2935–2964.
[2] C. Conley, Isolated Invariant Sets and the Morse Index, CBMS Lecture Notes, vol. 38, American Mathematical Society, Providence, RI, 1978.
[3] S. Day, Y. Hiraoka, K. Mischaikow, T. Ogawa, Rigorous numerics for global dynamics: a study of the Swift–Hohenberg equation, SIAM J. Appl. Dynam. Syst. 4 (2005) 1–31.
[4] Y. Hiraoka, Topological verification in infinite dimensional dynamical systems, Doctoral Dissertation, Osaka University, 2005.
[5] Y. Hiraoka, T. Ogawa, Rigorous numerics for localized patterns to the quintic Swift–Hohenberg equation, Jpn. J. Ind. Appl. Math. 22 (2005) 57–75.
[6] G. Iooss, M.C. Pérouéme, Perturbed homoclinic solutions in reversible 1:1 resonance vector fields, J. Differential Equations 102 (1993) 62–88.
[7] H.B. Keller, Lectures on Numerical Methods in Bifurcation Problems, Springer, Berlin, Notes by A.K. Nandakumaran, M. Ramaswamy, Indian Institute of Science, Bangalore, 1987.
[8] A. Mielke, G. Schneider, Attractors for modulation equations on unbounded domains—existence and comparison, Nonlinearity 8 (1995) 743–768.
[9] S. Orszag, Numerical simulation of incompressible flows within simple boundaries, I. Galerkin (spectral) representations, Stud. Appl. Math. 50 (1971) 293–327.
[10] H. Sakaguchi, H.R. Brand, Stable localised solutions of arbitrary length for the quintic Swift–Hohenberg equation, Physica D 97 (1996) 274–285.
[11] D. Salamon, Connected simple systems and the Conley index of isolated invariant sets, Trans. Amer. Math. Soc. 291 (1985) 1–41.
[12] J.B. Swift, P.C. Hohenberg, Hydrodynamic fluctuations at the convective instability, Phys. Rev. A 15 (1977) 319–328.
[13] P.D. Woods, A.R. Champneys, Heteroclinic tangles and homoclinic snaking in the unfolding of a degenerate reversible Hamiltonian–Hopf bifurcation, Physica D 129 (1999) 147–170.
[14] P. Zgliczyński, K. Mischaikow, Rigorous numerics for partial differential equations: the Kuramoto–Sivashinsky equation, Found. Comput. Math. 1 (2001) 255–288.